Adaptive Density Estimation on the Circle by Nearly-Tight Frames

Claudio Durastanti*

Abstract This work is concerned with the study of asymptotic properties of nonparametric density estimates in the framework of circular data. The estimation procedure here applied is based on wavelet thresholding methods: the wavelets used are the so-called Mexican needlets, which describe a nearly-tight frame on the circle. We study the asymptotic behaviour of the L^2-risk function for these estimates, in particular its adaptivity, proving that its rate of convergence is nearly optimal.

AMS 2010 Subject Classification - primary: 62G07; secondary: 62G20, 65T60, 62H11.

Key words: Density estimation, circular and directional data, thresholding, Mexican needlets, nearly-tight frames.

1 Introduction

In this work, we aim to study nonparametric estimation of a density function F based on directional data, sampled over the unit circle $\mathbb{S}^1$, by thresholding techniques, focussing in particular on the adaptivity for the associated L^2, the so-called mean integrated squared error. The estimators are built over a wavelet system, namely the Mexican needlets, which describes a nearly tight frame on $\mathbb{S}^1$ and characterized by strong localization in the spatial domain. Directional data over $\mathbb{S}^1$ can be viewed as angles measured with respect to a fixed starting point, the origin, and a fixed positive direction. They can be described as a set of points $\{X_i, i = 1, ..., n\}$, lying on

Claudio Durastanti
Ruhr Universität, D44780, Bochum
e-mail: claudio.durastanti@gmail.com

* This research is supported by DFG Grant n.2131

the circumference of $\mathbb{S}^1$: for this reason, they are also called circular data. Circular data are characterized by the 2π-periodicity, which has led to the development of a huge set of circular statistical methods, independently from the standard real-line statistics. These investigations can be motivated also in view of the large number of applications in many different fields, as for instance geophysics, oceanography and engineering. The textbooks [29, 14] can provide a complete overview on this topic and further technical details (see also [4, 31]), while some applications of interest can be found in [2, 5, 21, 22, 34].

1.1 Overview

In the recent years, the literature concerning density estimation problems is becoming more and more abundant: in particular, we are referring to the study of the adaptivity results for L^2-risks in the nonparametric framework. Consider a function F belonging to some scale of classes $\mathscr{F}_\alpha$, called nonparametric regularity class of functions and depending on a set of parameters $\alpha \in A$, and its estimator $\widehat{F}$: this estimator is said to be adaptive for the L^2-risk and for $\mathscr{F}_\alpha$, if for any $\alpha \in A$ there exists a constant c_α such that

$$\left\|\widehat{F}-F\right\|^2_{L^2\left(\mathbb{S}^1\right)} \leq c_\alpha R_n\left(\widehat{F},\mathscr{F}_\alpha\right),$$

where n is the number of sampled data and $R_n\left(\widehat{F},\mathscr{F}_\alpha\right)$ is, loosely speaking, the worst possible performance over $\mathscr{F}_\alpha$. it is said to be minimax if $R_n\left(\widehat{F},\mathscr{F}_\alpha\right) = \inf_F \sup_{\widehat{F}\in\mathscr{F}_\alpha} \left\|\widehat{F}-F\right\|^2_{L^2\left(\mathbb{S}^1\right)}$, where F ranges over all measurable functions of the observations $\{X_i, i = 1,\ldots,n\}$.
Nonparametric minimax estimation of unknown densities or regression functions was presented in the seminal paper [6], see also [7]: in this work, optimal minimax rates of convergence of the L^2-risk were obtained by nonlinear wavelet estimators based on thresholding techniques. Since then on, many applications were developed not only in Euclidean spaces but also in more general manifolds: we suggest as textbook reference [20]. As far as data on the unit q-dimensional sphere $\mathbb{S}^q$ are concerned, many of those researches have been developed by using the constructions of second-generation wavelets on $\mathbb{S}^q$ named spherical needlets. The spherical needlets, introduced in the literature by [26, 27], feature properties fundamental to attain the minimax optimal rates of convergence of the estimates, such as their concentration in both Fourier and space domains: density estimation of directional data on $\mathbb{S}^q$ was presented in [3], the analysis of nonparametric regression on sections of spin fiber bundles on $\mathbb{S}^2$ by the means of spin needlets was proposed in [8] and, finally, nonparametric regression estimators on the sphere based respectively on needlet block and global thresholding were studied in [11] and [13].

1.2 Motivations and comparisons with standard needlets

The main result here established concerns nearly-optimal rates of convergence for the L^2-risk of nonparametric density estimation based on wavelet coefficients on $\mathbb{S}^1$. The wavelets considered are the so-called Mexican needlets, introduced on general compact manifolds in [15, 16, 17, 18], see also [19, 28]. These wavelets are known to enjoy very good localization properties in the real domain, as described in details below in Section 2 (see also [10]), while their support is not bounded in the harmonic domain, on the contrary of standard needlets. Furthermore, while standard needlets are built by using a set of exact cubature points and weights (cfr. [26]), Mexican needlets are built over a set of points satisfying weaker restrictions (see [17] and Theorem 1 below). Indeed, Mexican needlets can be built over any partition over their spatial support with area monotonically decreasing with the resolution level. In this sense, statistical techniques adopting Mexican needlets are more immediately applicable for computational developing: some examples of their practical applications in the field of statistics can be found, for instance, in [9, 12, 23, 25, 30]. On the other hand, Mexican needlets lack an exact reconstruction formula, so that the corresponding density estimators are biased. The main purpose of this work is to show that thresholding procedures built on Mexican needlets behave asymptotically as those constructed with standard needlets (cfr. [3]), on the other hand offering advantages both from the practical and the theoretical points of view, such as the easier construction of the wavelets over partitions on $\mathbb{S}^1$ and the stronger localization properties; their bias is proved to be asymptotically negligible (see Theorem 3 below and numerical evidence in Section 5.

1.3 Statement of the main result

Given a set of i.i.d. circular data $\{X_i, i=1,...,n\}$, distributed over $\mathbb{S}^1$ with density F, and the set of circular Mexican needlets, $\left\{\psi_{jq;s}(\theta), \theta \in \mathbb{S}^1\right\}$, whose definition and main properties will be given below in Subsection 2.1, a threshold wavelet estimator $\widehat{F}$ for the density function is given by

$$\widehat{F}(\theta) = \sum_{j=J_0}^{J_n} \sum_{q=1}^{Q_j} \zeta_{jq}(\tau_n) \widehat{\beta}_{jq;sK} \psi_{jq;sK}(\theta) \ , \theta \in \mathbb{S}^1,$$

where $\zeta_{jq}(\tau_n)$ denotes the threshold, $\widehat{\beta}_{jq;sK}$ the unbiased estimator of the wavelet coefficient corresponding to $\psi_{jq;sK}(\theta)$, K is the cut-off frequency; further details can be found in Section 3. We choose Besov spaces, labelled by $\mathscr{B}^r_{m,t}$, as nonparametric regression class of functions, (cfr. Subsection 2.2), so that Theorem 2 will prove that

$$\sup_{F\in\mathscr{B}^r_{m,t}} \mathbb{E}\left[\left\|\widehat{F}-F\right\|^2_{L^2(\mathbb{S}^1)}\right] = O_n\left(\log n\left(\frac{n}{\log n}\right)^{-\frac{2r}{2r+1}}\right), \tag{1}$$

where r is one of the smoothness parameters characterizing the Besov space. Observe that the results here obtained are consistent with the ones already existing literature, cfr. for instance [3, 6, 7, 20]. We stress again that the estimator $\widehat{F}$ is characterized by a bias due to the lack of an exact reconstruction formula: the nearly-tightness of $\left\{\psi_{jq,s}(\theta), \theta\in\mathbb{S}^1\right\}$ assures the bias to be negligible with respect the rate of convergence on the left hand of (1), since it is controlled by some parameters depending on the number of observations n. All the details can be found in Theorem 3. We stress again that the study of the asymptotic behaviour of the bias is one of the most relevant results attained in this paper, because it represents the main difference between density estimates here defined and the ones built on standard needlets (see again [3]).

1.4 Plan of the paper

Section 2 introduces the circular Mexican needlets, their main properties and a quick overview on circular Besov spaces. In Section 3 we describe the nonparametric density estimates built on circular Mexican needlets, while Section 4 describes our main results: Theorem 2, concerns adaptivity of the threshold density estimator $\widehat{F}$ and Theorem 3 exploits the upper bound for the bias of $\widehat{F}$. Section 5 provides some numerical evidence, while in Section 6 we collect all the auxiliary results related to the two main theorems and some ancillary results on circular Mexican needlets.

2 Nearly-tight frames on the circle

This section will provide details concerning the construction and properties of Mexican needlet frames over $\mathbb{S}^1$ and the definition of circular Besov spaces in terms of their approximation properties.

2.1 Harmonic analysis and circular Mexican needlets

In this subsection we will describe some results, already well-known in the literature, related to Fourier analysis and the construction of the Mexican needlets over the unit circle $\mathbb{S}^1$. More details on Fourier analysis can be found, for instance, on the textbook [32], while Mexican needlets and, more in general, nearly-tight frames over compact manifolds were introduced in the literature in [15, 16, 17, 18], see also [19, 28]. Furthermore, we present also a simplified statement of the localiza-

tion property in the spatial domain for Mexican needlets over $\mathbb{S}^1$, described more extensively in Lemma 1 (see also [10, 17]).
Let us denote by $L^2\left(\mathbb{S}^1\right) \equiv L^2\left(\mathbb{S}^1, d\rho\right)$ the space of square integrable functions over the circle with respect to the Lebesgue measure $\rho\left(d\theta\right) = \left(2\pi\right)^{-1} d\theta$, on which we define the inner product as follows: for $f, g \in L^2\left(\mathbb{S}^1, d\rho\right)$

$$\left\langle f, g\right\rangle \equiv \left\langle f, g\right\rangle_{L^2\left(\mathbb{S}^1\right)} = \int_{\mathbb{S}^1} f\left(\theta\right) \overline{g\left(\theta\right)} \rho\left(d\theta\right),$$

As well known in the literature, the set $\left\{u_k\left(\theta\right), \theta \in \mathbb{S}^1, k \in \mathbb{Z}\right\}$, $u_k\left(x\right) = \exp\left(ik\theta\right)$, describes an orthonormal basis over $\mathbb{S}^1$, whereas the Fourier transform is given by

$$a_k = \left\langle f, u_k\right\rangle_{L^2\left(\mathbb{S}^1\right)} = \frac{1}{2\pi} \int_0^{2\pi} f\left(\theta\right) \overline{u_k\left(\theta\right)} d\theta,$$

and the corresponding Fourier inversion is given by

$$f\left(\theta\right) = \sum_{k \in \mathbb{Z}} a_k u_k\left(\theta\right)\ ,\ \theta \in \mathbb{S}^1. \tag{2}$$

Furthermore, $\left\{u_k\left(\theta\right), \theta \in \mathbb{S}^1, k \in \mathbb{Z}\right\}$ can be viewed as the eigenfunctions of the circular Laplacian Δ corresponding to eigenvalues $-k^2$ (for more details, see for instance [24]). For $F \in L^2\left(\mathbb{S}^1\right)$, the quantity γ_k is given by

$$\gamma_k := \left|a_k\right|^2, \tag{3}$$

so that

$$\sum_{k \in \mathbb{Z}} \gamma_k = \sum_{k \in \mathbb{Z}} \left|a_k\right|^2 = \left\|F\right\|^2_{L^2\left(\mathbb{S}^1\right)}.$$

Remark 1. Since $\left\|F\right\|^2_{L^2\left(\mathbb{S}^1\right)} < \infty$, the sum $\sum_{k \in \mathbb{Z}} \gamma_k$ has to converge, therefore

$$\lim_{|k| \to \infty} \gamma_k = 0,$$
$$\lim_{|k| \to \infty} \left|a_k\right| = 0.$$

Let us now introduce the Mexican needlet system. Let the weight function $w_s : \mathbb{R} \mapsto \mathbb{R}_+$ be given by

$$w_s\left(x\right) := x^s \exp\left(-x\right)\ ,\ x \in \mathbb{R}, \tag{4}$$

so that, from the Calderon formula and for $t \in \mathbb{R}_+$, it holds that

$$e_s := \int_0^{\infty} \left|w_s\left(tx\right)\right|^2 \frac{dx}{x} = \frac{\Gamma\left(2s\right)}{2^{2s}},$$

while (see [17]) using the Daubechies' criterion leads us to

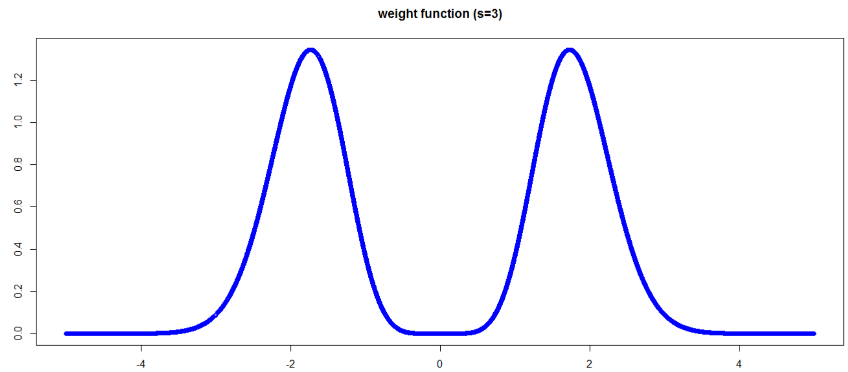

Fig. 1 the weight function $w_s\left(x^2\right)$ for $s=3$.

$$\Lambda_{B,s} m_B \leq \sum_{j=-\infty}^{\infty} \left| w_s\left(tB^{-2j}\right)\right|^2 \leq \Lambda_{B,s} M_B,$$

where, for the scale parameter $B>1$,

$$\begin{aligned}
\Lambda_{B,s} &= e_s\left(2\log B\right)^{-1},\\
M_B &= \left(1+O_B\left(|B-1|^2\log|B-1|\right)\right),\\
m_B &= \left(1-O_B\left(|B-1|^2\log|B-1|\right)\right).
\end{aligned}$$

For any resolution level $j\in\mathbb{Z}$, let $\left\{E_{jq}\right\}$, $q=1,\ldots,Q_j$ be a partition of $\mathbb{S}^1$, such that $E_{jq_1}\cap E_{jq_2}=\varnothing$ for $q_1\neq q_2$. Any E_{jq} is characterized by the couple (λ_{jq},x_{jq}): $\lambda_{jq}=\rho\left(E_{jq}\right)$ describes the length of E_{jq}, while $x_{jq}\in E_{jq}$ is a point belonging to E_{jq}. For the sake of simplicity, we can think to x_{jq} as the midpoint of the segment of arc E_{jq}. Fixed now the shape parameter $s\in\mathbb{N}$ and the scale parameter $B>1$, the circular Mexican needlet $\psi_{jq;s}:\mathbb{S}^1\mapsto\mathbb{C}$ is given by

$$\begin{aligned}
\psi_{jq;s}\left(\theta\right) &:= \sqrt{\lambda_{jq}}\sum_{k=-\infty}^{\infty} w_s\left(\left(B^{-j}k\right)^2\right)\overline{u_k\left(x_{jq}\right)}u_k\left(\theta\right)\\
&= \sqrt{\lambda_{jq}}\sum_{k=-\infty}^{\infty} w_s\left(\left(B^{-j}k\right)^2\right)\exp\left(ik\left(\theta-x_{jq}\right)\right),\ \theta\in\mathbb{S}^1. \qquad (5)
\end{aligned}$$

For any $F\in L^2\left(\mathbb{S}^1\right)$, the needlet coefficient $\beta_{jq;s}\in\mathbb{C}$ corresponding to $\psi_{jq;s}$ is given by

$$\beta_{jq;s} := \left\langle F,\psi_{jq;s}\right\rangle_{L^2\left(\mathbb{S}^1\right)}. \qquad (6)$$

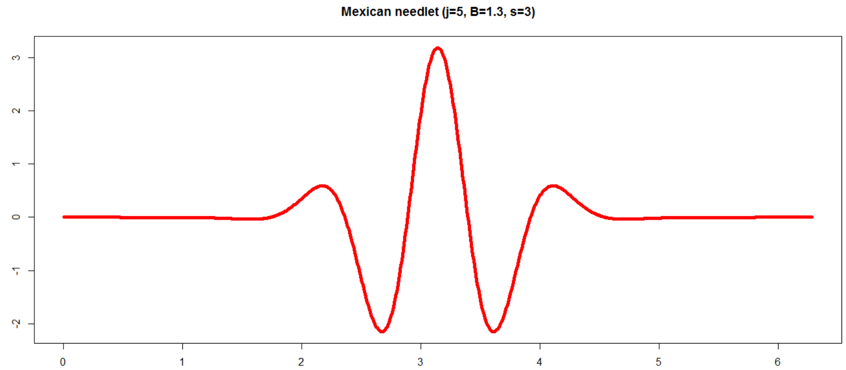

Fig. 2 The Mexican Needlet with $s=3, B=1.3, j=5$ centered on the point $x_{jq}=\pi$.

The next result, here properly fitted for $\mathbb{S}^1$, was originally proposed as Theorem 1.1 in [17]: it proves that the Mexican needlet framework describes a nearly-tight frame on the manifold. We recall that a set of functions $\{e_i, i \geq 1\}$ defined over a manifold M is a frame if there exist $c_1, c_2 > 0$ so that, for any $F \in L^2(M)$,

$$c_1 \|F\|^2_{L^2(M)} \leq \sum_i \left| \langle F, e_i \rangle_{L^2(M)} \right|^2 \leq c_2 \|F\|^2_{L^2(M)} .$$

A frame is said to be tight if $c_1 = c_2$. An example of a tight frame over the d-dimensional sphere $\mathbb{S}^d$ is given by the standard needlets, introduced in the literature in [26, 27]. A frame is nearly-tight if $c_2/c_1 \simeq 1+\varepsilon$, where ε is close to 0, cfr. [16].

Theorem 1. ***(Nearly-tightness of the Mexican needlets frame - Th. 1.1 in [17])*** *Fixing $B>1$ and $c_0, \delta_0 > 0$ sufficiently small, there exists a constant C_0 as follows:*

- *for $0<\eta<1$, suppose that for each $j \in \mathbb{Z}$, there exists a set of measurable sets $\left\{E_{jq}, q=1,\ldots,Q_j\right\}$, with $\lambda_{jq} = \mu\left(E_{jq}\right)$, where:*
  - *$\lambda_{jq} \leq \eta B^{-j}$;*
  - *for each j with $\eta B^{-j} < \delta_0$, $\lambda_{jq} \geq c_0 \left(\eta B^{-j}\right)$ for $q=1,\ldots,Q_j$;*
- *it holds that*

$$\left(\Lambda_{B,s} m_B - C_0 \eta\right) \|F\|^2_{L^2\left(\mathbb{S}^1\right)} \leq \sum_{j=-\infty}^{\infty} \sum_{q=1}^{Q_j} \left|\beta_{jq;s}\right|^2 \leq \left(\Lambda_{B,s} M_B + C_0 \eta\right) \|F\|^2_{L^2\left(\mathbb{S}^1\right)} .$$

If $\left(\Lambda_{B,s} m_B - C_0 \eta\right) > 0$, then $\left\{\psi_{jq;s}\right\}$ is a nearly tight frame, since

$$\frac{\left(\Lambda_{B,s} M_B + C_0 \eta\right)}{\left(\Lambda_{B,s} m_B - C_0 \eta\right)} \sim \frac{M_B}{m_B} = 1 + O_B\left(|B-1|^2 \log|B-1|\right).$$

Mexican needlets can be thought as an alternative approach to the standard needlets, proposed in [26, 27], see also [3, 24], in views of their stronger localization property in the real domain. Standard needlets feature a quasi-exponential localization property in the spatial range, while the weight function w_s leads to a full-exponential localization in the real space as proved below in Lemma 1 (cfr. [10, 17]). As far as the frequency domain is concerned, while spherical needlets lie on compact support (see again [26, 27]), each Mexican needlet has to take in account the whole frequency range. This issue is partially compensated by the structure itself of the function w_s, exponentially localized around a dominant term in the frequency domain and, therefore, consistently different from zero only on limited set of frequencies. For our purposes and in order to respect the conditions in Theorem 1, we impose the following

Condition 1. . *Let $\psi_{jk;s}(\theta)$ and $\beta_{jk;s}$ be given respectively by* (5) *and* (6). *We have that, for $j > 0$*

$$Q_j \approx \eta^{-1} B^j \ , \ \lambda_{jq} \approx \eta B^{-j},$$

so that

$$\psi_{jq;s}(\theta) \approx \eta^{\frac{1}{2}} B^{-\frac{j}{2}} \sum_{k=-\infty}^{\infty} w_s\left(\left(B^{-j}k\right)^2\right) \exp\left(ik\left(\theta - x_{jq}\right)\right) \ , \ \theta \in \mathbb{S}^1. \tag{7}$$

Furthermore, we choose $J_0 < -\log_B \sqrt{s}$ and fix δ_0 such that $\delta_0 \leq \eta B^{-J_0}$. Hence, we have, for $j < J_0$,

$$Q_j = 1, \ \lambda_{jq} = 2\pi. \tag{8}$$

Mexican needlets are characterized by the following localization property, proven in Lemma 1:

$$\left|\psi_{jk;s}(\theta)\right| \leq \sqrt{\lambda_{jk}} c_s B^j \exp\left(-\left(\frac{B^j\left(\theta - x_{jk}\right)}{2}\right)^2\right)\left(1 + \left(\frac{B^j\left(\theta - x_{jk}\right)}{2}\right)^{2s}\right).$$

From the localization property, it follows a bound rule on the norms: there exist $\widetilde{c_p}, \widetilde{C_p} > 0$ such that

$$\widetilde{c_p} B^{j\left(\frac{p}{2}-1\right)} \eta^{\frac{p}{2}} \leq \left\|\psi_{jq;s}\right\|^p_{L^p\left(\mathbb{S}^1\right)} \leq \widetilde{C_p} \eta^{\frac{p}{2}} B^{j\left(\frac{p}{2}-1\right)}. \tag{9}$$

The proof, totally analogous to the case of standard needlets (see [27]), is here omitted.

Remark 2. The choice of (8) is justified as follows. First of all, observe that, for any $j < J_0$, λ_{jq} still satisfies Theorem 1. Furthermore, when j is negative, the B^{-j} grows to infinity, hence there exists some $J' < 0$ such that $\delta_0 \leq \eta B^{-J'}$. It implies that the λ_{jq} has to be smaller than a quantity bigger than $4\pi = \rho\left(\mathbb{S}^1\right)$, corresponding to the case $E_{jq} \equiv \mathbb{S}^1$, which leads to $Q_j = 1$, so that we have that $Q_j \lambda_{jk} \approx 1$. As far as the choice of J_0 is concerned, if $J_0 < -\log_B \sqrt{s}$, it means that, for any k, $\left|kB^{-J_0}\right| > s$,

and therefore $w_s\left(\left(kB^{-J_0}\right)^2\right) < w_s(s) = \max_{r\in\mathbb{R}} w_s(r)$. As consequence, taking into account Lemma 2, we have that for any k, $\chi_{s,B,J_0}\left(k^2\right) << 2^{-2s}\Gamma(2s) = e_s$.

Remark 3. While in [15, 16, 17, 18] the Mexican needlets are defined as $\psi_{j'q}(\theta) \equiv \psi_{-jq}(\theta)$, $\theta \in \mathbb{S}^1$. We use this notation to uniform this work to the already existing literature on the field of statistics based on needlet-like framework.

2.2 Besov spaces on the circle

In this subsection, we will recall some of the results proposed in [18] (see also [20, 27]) on Besov spaces, in terms of their approximation properties. More in details, let Π_r be the space of polynomials of degree r: we start by looking for the infimum of the $L^p\left(\mathbb{S}^1\right)$-distance between a function $f : \mathbb{S}^1 \mapsto \mathbb{R}$ and the space Π_r:

$$G_r(f,P) = \inf_{P\in\Pi_r} \|f-P\|_{L^p\left(\mathbb{S}^1\right)}.$$

Following, for instance, [3, 8, 18, 27], let $F \in \mathscr{B}^r_{m,t}$, if and only if both the following conditions hold:

$$(i)\ F\in L^m\left(\mathbb{S}^1\right)\ ,\ (ii)\left(\sum_u \left(u^s G_u(f,P)\right)^{\frac{r}{u}}\right)^{\frac{1}{r}},$$

or, equivalently,

$$(i)\ F\in L^m\left(\mathbb{S}^1\right)\ ,\ (ii)\left(\sum_j \left(B^{-jr} G_{B^j}(f,P)\right)^{q}\right)^{\frac{1}{q}}.$$

As shown in [18], see also [3], it holds that, for $1 \leq m \leq \infty$, $r > 0$, $0 \leq t \leq \infty$, $f \in \mathscr{B}^r_{m,t}$ if and only if

$$\left(\sum_{q=1}^{Q_j} \left|\beta_{jq;s}\right|^m \left\|\psi_{jk;s}\right\|^m_{L^m\left(\mathbb{S}^1\right)}\right)^{\frac{1}{m}} < B^{-jr}\delta_j,\ \delta_j \in \ell_r.$$

In what follows, we will make extensive use of this inequality with $m = 2$:

$$\left(\sum_{q=1}^{Q_j} \left|\beta_{jq;s}\right|^2 \left\|\psi_{jk;s}\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right)^{\frac{1}{2}} \leq \left(\widetilde{C_2}\eta \sum_{q=1}^{Q_j} \left|\beta_{jq;s}\right|^2\right)^{\frac{1}{2}} < B^{-jr}\delta_j,\ \delta_j \in \ell_r. \quad (10)$$

Further details on Besov spaces can be found in [27] and in the textbook [20].

3 The density estimation procedure

In this section, we will introduce a thresholding density function estimator on the circle based on the Mexican needlet coefficients. As already mentioned, thresholding techniques were introduced in the literature by D. Donoho and I. Johnstone in [6], to be later successfully applied in several research topics: for an exhaustive overview and details we suggest the textbooks [20] and [33]. Consider a set of random directional observations $\left\{X_i \in \mathbb{S}^1 : i = 1, ..., n\right\}$ with common distribution $\nu(\theta) = F(\theta)\, d\theta$ and let us introduce the threshold function $\zeta_{jq}(\tau_n) := \mathbb{1}_{\left\{\left|\beta_{jq;s}\right| \geq \kappa\tau_n\right\}}$, where κ is a real-valued positive constant to be chosen to set the size of the threshold (cfr. [3]). The coefficient estimator is given by

$$\widehat{\beta}_{jq;s} := \frac{1}{n}\sum_{i=1}^{n} \overline{\Psi}_{jq;sK}(X_i),$$

which is unbiased, i. e.

$$\mathbb{E}\left[\widehat{\beta}_{jqK;s}\right] = \int_{\mathbb{S}^1} \overline{\Psi}_{jq;sK} F(\theta)\, d\theta = \beta_{jqK;s}.$$

Consequently, the thresholding density estimator is given by

$$\widehat{F}(\theta) = \sum_{j=J_0}^{J_n}\sum_{q=1}^{Q_j} \zeta_{jq}(\tau_n)\, \widehat{\beta}_{jqK;s} \psi_{jq;sK}(\theta)\ ,\ \theta \in \mathbb{S}^1, \tag{11}$$

where J_n and K_n represent respectively the truncation resolution level and the cut-off frequency. The truncation level is chosen so that $B^{J_n} = \sqrt{\frac{n}{\log n}}$, as usual in the literature (see for instance [3, 8]), while the cut-off frequency is fixed so that $K_n = \sqrt{\frac{n}{\log n}}$, . The other tuning parameters of the Mexican needlet estimator to be considered are:

- the threshold constant κ, whose evaluation is given in the Section 6 of [3];
- the scaling factor τ_n, depending on the sample size, chosen, as usual in the literature, as $\tau_n = (\log n/n)^{1/2}$;
- the pixel-parameter $\eta_n = \eta$, chosen so that $\eta_n = O_n\left(n^{-\frac{2}{3}}\right)$.

We will present our main result concerning Mexican thresholding density estimation in the next Theorem. For the embeddings featured by the Besov spaces, as in [3], the condition $r > \frac{1}{m}$ implies that $F \in \mathscr{B}^r_{m,t} \subset \mathscr{B}^{r-\frac{1}{m}}_{\infty,t}$, so that F is continuous.

Theorem 2. *For $1 \leq m = t < 2$, $r > \frac{1}{m}$, there exists some constant $C_0 = C_0(m,r)$ such that*

$$\sup_{F \in \mathscr{B}^r_{m,t}} \mathbb{E}\left[\left\|\widehat{F} - F\right\|^2_{L^2(\mathbb{S}^1)}\right] \leq C_0 \log n \left(\frac{n}{\log n}\right)^{-\frac{2r}{2r+1}}. \tag{12}$$

Remark 4. To attain optimality, it should be necessary to show also that

$$\sup_{F\in\mathscr{B}^r_{m,t}} \mathbb{E}\left[\left\|\widehat{F}-F\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right] \geq C_* \left(\frac{n}{\log n}\right)^{-\frac{2r}{2r+1}} .$$

This lower bound is entirely analogous to the standard needlet case in [3], Theorem 11, and therefore its proof is here omitted.

4 Proof of Theorem 2

In this section we will provide a proof for Theorem 2 based on the main guidelines described by D. Donoho and I. Johnstone in [6], cfr. also [3] and the textbooks [20, 33]. The procedure illustrated by [3, 6, 7] fits perfectly for tight wavelet systems, which feature an exact reconstruction formula. As already discussed in Subsection 2.1, Mexican needlet are not characterized by tightness, hence the bias term appearing in the study of (12) will also take into account addends due the (deterministic) error raising when we approximate a function with its wavelet expansion. The decay of these terms will depend on the choice of the pixel-parameter η_n, on one hand, and of J_n and K_n on the other hand. We will start by developing an upper bound for the bias term, which represents the main difference between the estimation procedure here discussed and the one based on standard needlet frames.

4.1 The bias: the construction and the upper bound

We recall from [17] the so-called summation operator S, leading to the *summation formula*. The summation formula can be viewed as the equivalent in the Mexican needlet framework of the reconstruction formula in the standard needlet case (see for instance [26, 24]): for any $F \in L^2\left(\mathbb{S}^1\right)$, let the *summation operator* $S\left[F\right]_s$ be given by

$$S\left[F\right]_s(\theta) := \sum_{j=J_0}^{\infty}\sum_{q=1}^{Q_j} \beta_{jq;s}\psi_{jq;s}(\theta) \ , \theta \in \mathbb{S}^1. \tag{13}$$

The goal of this subsection is also to estimate which terms in the sum above are so small that they can be neglected. We will fix a cut-off frequency K, to compensate the lack of a compact support in the harmonic domain typical of standard needlets (see [26]), to define the truncated Mexican needlet, and a truncation resolution level J. Theorem 3 will exploit an upper bound, depending on s, J, K and η, between (13) and the truncated summation operator defined below. Observe that these results are general and not related to the specificity of the estimation problem: in this sense, when the label n in K and J is omitted, we intend that the claimed result holds in general.

First of all, given $K \in \mathbb{N}$, the *truncated Mexican needlet* $\psi_{jq;sK}$ is given by

$$\psi_{jq;sK}(\theta) := \sqrt{\lambda_{jq}} \sum_{|k| \leq K} w_s\left(\left(kB^{-j}\right)^2\right) \overline{u_k(\xi_{jq})} u_k(\theta)\ ,\ \theta \in \mathbb{S}^1,\ \xi_{jq} \in E_{jq}\ ,$$

and the corresponding *truncated needlet coefficient* $\beta_{jq;sK}$ is defined

$$\beta_{jq;sK} := \left\langle F, \psi_{jq;sK} \right\rangle_{L^2\left(\mathbb{S}^1\right)}\ .$$

Loosely speaking, fixed K, $\psi_{jk;sK}(\cdot)$ is the Mexican needlet where all the elements out of the support $[-K, K]$ are not taken into account. The *truncated summation operator* $S[F]_{s,K,J}$ is therefore given by

$$S[F]_{s,K,J}(\theta) := \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \beta_{jq;sK} \psi_{jq;sK}(\theta)\ ,\ \theta \in \mathbb{S}^1. \tag{14}$$

Remark 5. Following Remark 2, we will truncate in (14) all the negative resolution levels $j < J_0$.

Let the bias $R_{s,K,J,\eta}$ be given by

$$R_{s,K,J,\eta} := \left\| S[F]_s - S[F]_{s,K,J} \right\|_{L^2\left(\mathbb{S}^1\right)}\ ; \tag{15}$$

An upper bound for $R_{s,K,J}$ is explicitly provided in the next Theorem.

Theorem 3. *Let $R_{s,K,J}$ be given by* (15)*. Then, there exist $C_1, C_2, C_3 > 0$ such that*

$$R_{s,K,J} \leq C_1 B^{-rJ} + C_2 J^{\frac{1}{2}} K^{2s-\frac{1}{2}} \exp\left(-K^2\right) B^{-\left(r+2s-\frac{1}{2}\right)J}$$
$$+ C_3 B^{(1-2s)J} J^{\frac{1}{2}} K^{s-\frac{1}{4}} e^{-2K^2} \left(\sum_{|k|>K} \gamma_k\right)^{\frac{1}{2}}$$

Proof. Using the Minkowski inequality, we have

$$\left\| S[F]_s - S[F]_{s,K,J} \right\|_{L^p\left(\mathbb{S}^1\right)} \leq I_1 + I_2 + I_3,$$

where

$$I_1 := \left\| \sum_{j=J_0}^{\infty} \sum_{q=1}^{Q_j} \beta_{jq;s} \psi_{jq;s} - \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \beta_{jq;s} \psi_{jq;s} \right\|_{L^2(\mathbb{S}^1)} ;$$
$$I_2 := \left\| \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \beta_{jq;s} \psi_{jq;s} - \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \beta_{jq;s} \psi_{jq;s,K} \right\|_{L^2(\mathbb{S}^1)} ;$$
$$I_3 := \left\| \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \beta_{jq;s} \psi_{jq;sK} - \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \beta_{jq;sK} \psi_{jq;sK} \right\|_{L^2(\mathbb{S}^1)} .$$

Observe that while I_1 describes the bias due to the truncation of the resolution levels belonging to (J,∞), I_2 and I_3 depend strictly on the choice of the cut-off frequency K, due to the approximation error due to approximate Mexican needlets by the corresponding truncated ones (the former), and Mexican needlet coefficients by the corresponding truncated ones (the latter). According to Lemma 6, we get

$$I_1 \leq C_{1,1} B^{-rJ}.$$

As far as I_2 is concerned, from Lemma 7, we obtain

$$I_2 \leq C_2 J^{\frac{1}{2}} K^{2s-\frac{1}{2}} \exp\left(-K^2\right) B^{-\left(r+2s-\frac{1}{2}\right)J} .$$

Finally, from Lemma 8, it holds that

$$I_3 \leq C_3 B^{(1-2s)J} J^{\frac{1}{2}} K^{s-\frac{1}{4}} e^{-2K^2} \left(\sum_{|k|>K} \gamma_k \right)^{\frac{1}{2}} ,$$

as claimed.

Remark 6. An analogous result is obtained in Theorem 2.5 in [17] (see also Lemma 2.3 in [16]). In these works, the authors use a generic weight function belonging to the Schwarz space and, moreover, the wavelets studied are defined over a general compact manifold. For this reason, the bound exploited in Theorem 3, using explicit bounds provided by w_s and by the basis $\{u_k\}$, is more precise.

4.2 Adaptivity of $\widehat{F}$ for the L^2-risk

Merging the results achieved in the previous subsection with the ones driven by the standard procedure in the case of nonparametric thresholding density estimation (see for instance [3]), we obtain the following proof.

Proof (Proof ot the Theorem 2). Observe that, for the triangular inequality, we have

$$\mathbb{E}\left[\left\|\widehat{F}-F\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right]=\mathbb{E}\left[\left\|\widehat{F}-S\left[F\right]_{s,K_n,J_n}+S\left[F\right]_{s,K_n,J_n}-S\left[F\right]_s+S\left[F\right]_s-F\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right]$$
$$\leq E_1+E_2+E_3\ ,$$

where

$$\begin{aligned}
E_1 &= \mathbb{E}\left[\left\|\widehat{F}-S\left[F\right]_{s,K_n,J_n}\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right]\ ;\\
E_2 &= R^2_{s,K_n,J_n}\ ;\\
E_3 &= \left\|S\left[F\right]_s-F\right\|^2_{L^2\left(\mathbb{S}^1\right)}\ .
\end{aligned}$$

As far as E_1 is concerned, the bound is established analogously to the one achieved in [3], hence here it is given just a sketch of this proof in Lemma 9 (see also [8]). Indeed, we have:

$$\begin{aligned}
E_1 &= \mathbb{E}\left[\left\|\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\left(\zeta_{jq}\left(\tau_n\right)\widehat{\beta}_{jq;sK_n}-\beta_{jq;sK_n}\right)\psi_{jq;sK_n}\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right]\\
&\leq \left(J_n+1\right)\sum_{j=0}^{J_n}\mathbb{E}\left[\left\|\sum_{q=1}^{Q_j}\left(\zeta_{jq}\left(\tau_n\right)\widehat{\beta}_{jq;sK_n}-\beta_{jq;sK_n}\right)\psi_{jq;sK_n}\right\|^2_{L^2\left(\mathbb{S}^1\right)}\right]\\
&\leq \left(J_n+1\right)\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\left\|\psi_{jq;sK_n}\right\|^2_{L^2\left(\mathbb{S}^1\right)}\mathbb{E}\left[\left|\zeta_{jq}\left(\tau_n\right)\widehat{\beta}_{jq;sK_n}-\beta_{jq;sK_n}\right|^2\right]\\
&\leq \left(J_n+1\right)\widetilde{C_2}\eta_n\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\mathbb{E}\left[\left|\zeta_{jq}\left(\tau_n\right)\widehat{\beta}_{jq;sK_n}-\beta_{jq;sK_n}\right|^2\right]\\
&\leq C_1J_n\left(E_{1,1}+E_{1,2}+E_{1,3}+E_{1,4}\right)\ ,
\end{aligned}$$

where

$$E_{1,1}=\eta_n\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\mathbb{E}\left[\left|\zeta_{jq}\left(\tau_n\right)\widehat{\beta}_{jq;sK_n}-\beta_{jq;sK_n}\right|^2\mathbb{1}_{\left\{\left|\widehat{\beta}_{jq;sK_n}\right|\geq\kappa\tau_n\right\}}\mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right|\geq\frac{\kappa\tau_n}{2}\right\}}\right]; \tag{16}$$

$$E_{1,2}=\eta_n\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\mathbb{E}\left[\left|\zeta_{jq}\left(\tau_n\right)\widehat{\beta}_{jq;sK_n}-\beta_{jq;sK_n}\right|^2\mathbb{1}_{\left\{\left|\widehat{\beta}_{jq;sK_n}\right|\geq\kappa\tau_n\right\}}\mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right|\leq\frac{\kappa\tau_n}{2}\right\}}\right]; \tag{17}$$

$$E_{1,3}=\eta_n\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\left|\beta_{jq;sK_n}\right|^2\mathbb{E}\left[\mathbb{1}_{\left\{\left|\widehat{\beta}_{jq;sK_n}\right|<\kappa\tau_n\right\}}\mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right|\geq2\kappa\tau_n\right\}}\right]; \tag{18}$$

$$E_{1,4}=\eta_n\sum_{j=0}^{J_n}\sum_{q=1}^{Q_j}\left|\beta_{jq;sK_n}\right|^2\mathbb{E}\left[\mathbb{1}_{\left\{\left|\widehat{\beta}_{jq;sK_n}\right|<\kappa\tau_n\right\}}\mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right|<2\kappa\tau_n\right\}}\right]. \tag{19}$$

Heuristically, the cross-terms $E_{1,2}$ and $E_{1,3}$ are bounded by means of fast decays of the probabilistic inequalities given in Lemma 10, while as far as $E_{1,1}$ and $E_{1,4}$ are concerned, their bounds will be exploited according to the tail properties of the Besov spaces: further details are in Lemma 9. From these considerations, it follows that

$$E_1 \leq C_1 \left(\frac{n}{\log n} \right)^{-\frac{2r}{2r+1}} .$$

As far as E_2 is concerned, from Theorem 3, it holds that

$$\begin{aligned} E_2 = \leq C_1 B^{-rJ_n} + C_2 J^{\frac{1}{2}} K_n^{2s-\frac{1}{2}} \exp\left(-K_n^2\right) B^{-\left(r+2s-\frac{1}{2}\right)J_n} \\ + C_3 B^{(1-2s)J_n} J_n^{\frac{1}{2}} K_n^{s-\frac{1}{4}} \exp\left(-2K_n^2\right) \left(\sum_{|k|>K_n} \gamma_k \right)^{\frac{1}{2}} \end{aligned}$$

Observe that

$$B^{-2rJ_n} = \left(\frac{n}{\log n} \right)^{-r} \leq \left(\frac{n}{\log n} \right)^{-\frac{2r}{2r+1}} ,$$

while

$$C_2 J_n^{\frac{1}{2}} K_n^{2s-\frac{1}{2}} \exp\left(-K_n^2\right) B^{-\left(r+2s-\frac{1}{2}\right)J_n} \leq \frac{n}{\log n}^{-\frac{2r}{2r+1}}$$

$$B^{(1-2s)J_n} J_n^{\frac{1}{2}} K_n^{s-\frac{1}{4}} e^{-2K_n^2} \leq \frac{n}{\log n}^{-\frac{2r}{2r+1}}$$

Finally, we have

$$\| S[F]_s - F \|^2_{L^2\left(\mathbb{S}^1\right)} \leq C_3 \eta_n ,$$

whence, for $r > 1$,

$$\eta_n \leq n^{-\frac{3}{4}} \leq \frac{n}{\log n}^{-\frac{3}{4}} \leq \frac{n}{\log n}^{-\frac{2r}{2r+1}} .$$

5 Numerical results

This section presents the results of some numerical experiments. Obviously, in the framework of finite sample situation, the asymptotic rate given in Theorem 2 has to be considered just as a prompt. In what follows, we have built an estimator (11) using the set to estimate $F(\theta) = (2\pi)^{-1} \exp\left((\theta - \pi)^2 / 2 \right)$ by using CRAN R. Some graphical evidence can be found in Figure 3. We will focus on two main points:

- the number of coefficients surviving to the thresholding procedure depending on κ and τ_n.

	n=8000			n=12000			
$j\backslash\kappa_0$	0.10	0.15	0.20	0.10	0.15	0.20	Tot
0	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	2	3
4	4	4	4	4	3	3	4
5	5	4	4	5	5	5	5
6	8	7	6	8	7	6	8
7	11	10	10	10	10	9	11
8	13	11	11	13	14	14	15
9	19	14	12	20	21	18	21
10	17	4	3	28	12	10	29
11	NA	NA	NA	7	2	0	40

Table 1 number of mexican needlet coefficients surviving thresholding for various values of n, j and κ_0.

	n=8000			n=12000		
κ_0	0.10	0.15	0.20	0.10	0.15	0.20
L^2-risk	0.481	0.468	0.451	0.458	0.432	0.331

Table 2 L^2-risk for various values of n and κ_0.

- the estimate of the L^2-risk function $\left\|\widehat{F}-F\right\|_{L^2\left(\mathbb{S}^1\right)}$ depending on the number of observations n.

In particular, following [3], we have chosen $\kappa = \kappa_0\sqrt{0.107}\sup_{\theta\in\mathbb{S}^1}\left|F\left(\theta\right)\right|$, with $\kappa_0 = 0.05, 0.1, 0.15, 0.2$, and $n = 8000, 12000$, leading to $K_{8000} = 30$, $K_{12000} = 36$, $J_{8000} = 10$, $J_{12000} = 11$, $t_{8000} = 0.0335$ and $t_{12000} = 0.028$. The Table 5 counts the number of coefficients survived to thresholding. A qualitative analysis confirms that: (i) as $n \to \infty$, t_n is decreasing so that the threshold is lower and more $\widehat{\beta}_{jq;s}$ survive to the thresholding procedure and (ii) if κ_0 increases, the number of surviving coefficients is smaller, expecially at higher resolution levels.
The Table 5 describes the estimates of the L^2-risks for any choice of κ_0 and any n. As expected, the L^2-risk is decreasing when n grows and it is increasing with respect to κ_0 (cfr. [3])

6 Auxiliary results

This section contains all the statements and the proofs of the auxiliary results used to prove Theorem 2 and Theorem 3.

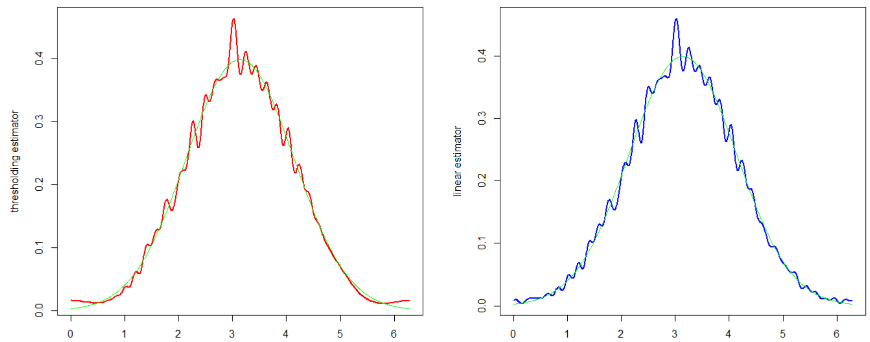

Fig. 3 Graphs of the thresholding estimator (on the left) and of the linear (not-thresholded) estimator (on the right) for $n = 12000$ and $s = 3$.

6.1 Properties and Inequalities for Mexican needlets

The first result here presented concerns the concentration property of the Mexican needlets in the real domain.

Lemma 1. *For every $\theta \in \mathbb{S}^1$, $s \geq 1$, there exists c_s such that:*

$$\left|\psi_{jq;s}(\theta)\right| \leq \sqrt{\lambda_{jq}} c_s B^j \exp\left(-\left(\frac{B^j(\theta - x_{jq})}{2}\right)^2\right)\left(1+\left(\frac{B^j(\theta - x_{jq})}{2}\right)^{2s}\right).$$

Furthermore, if $j \geq 0$, it holds that

$$\left|\psi_{jq;s}(\theta)\right| \leq c_s \eta B^{\frac{j}{2}} \exp\left(-\left(\frac{B^j(\theta - x_{jq})}{2}\right)^2\right)\left(1+\left(\frac{B^j(\theta - x_{jq})}{2}\right)^{2s}\right).$$

Proof. This proof follows strictly the one for standard needlets developed in [26] and the one for Mexican needlets on $\mathbb{S}^2$ in [10], see also [24]. First of all, from (4) observe that we can define a function $W_s : \mathbb{R} \mapsto \mathbb{R}_+$ such that $W_s(x) = w_s(x^2)$. We can therefore rewrite (5) as follows:

$$\psi_{jq;s}(\theta) = \sqrt{\lambda_{jq}} \sum_{k=-\infty}^{\infty} g_{B^j,\theta - x_{jq}}(k),$$

where

$$g_{B^j,\theta,x_{jq}}(u) := W_s\left(uB^{-j}\right)\exp\left(iu\left(\theta - x_{jq}\right)\right).$$

For the Poisson summation formula (see for instance [26, 10, 24]), we have that

$$\sum_{k=-\infty}^{\infty} g_{B^j,\theta,x_{jq}}(k) = \sum_{\nu=-\infty}^{\infty} \mathscr{F}\left[g_{B^j,\theta,x_{jq}}\right](2\pi\nu),$$

where the symbol $\mathscr{F}[g]$ denotes the Fourier transform of g. In our case, we have

$$\mathscr{F}\left[g_{B^j,\theta x_{jq}}\right](\omega) = \mathscr{F}\left[W_s\left(uB^{-j}\right)\right] * \mathscr{F}\left[\exp\left(iu\left(\theta - x_{jq}\right)\right)\right],$$

where the symbol $*$ denotes the convolution product. Standard calculations lead to

$$\mathscr{F}\left[W_s\left(uB^{-j}\right)\right] = \frac{(-1)^s}{\sqrt{2B^{-j}}} H_{2s}\left(\frac{\omega}{2B^{-j}}\right) \exp\left(-\left(\frac{\omega}{2B^{-j}}\right)^2\right);$$
$$\mathscr{F}\left[\exp\left(iu\left(\theta - x_{jq}\right)\right)\right] = \sqrt{2\pi}\,\delta\left(\left(\theta - x_{jq}\right) - \omega\right).$$

Hence we get

$$\mathscr{F}\left[g_{B^j,\theta,x_{jq}}\right](\omega) = \frac{(-1)^s\sqrt{\pi}}{B^{-j}} H_{2s}\left(\frac{\left(\theta - x_{jq}\right) - \omega}{2B^{-j}}\right) \exp\left(-\left(\frac{\left(\theta - x_{jq}\right) - \omega}{2B^{-j}}\right)^2\right).$$

Following Proposition 2 in [10], we have that

$$\sum_{\nu=-\infty}^{\infty} \mathscr{F}\left[\widehat{g}_{B^j,\theta-x_{jq}}\right](2\pi\nu) \leq C_{2s} B^j \exp\left(-\left(\frac{B^j\left(\theta - x_{jq}\right)}{2}\right)^2\right) H_{2s}\left(\frac{B^j\left(\theta - x_{jq}\right)}{2}\right).$$

It can be easily proved that

$$H_{2s}\left(\frac{B^j\left(\theta - x_{jq}\right)}{2}\right) \approx \left(1 + \left(\frac{B^j\left(\theta - x_{jq}\right)}{2}\right)^{2s}\right),$$

see for instance [11]. Straightforward calculations lead to the claimed result.

The next results will be pivotal to truncate negative resolution levels and the frequencies $k : |k| > K$ in the summation formula.

Lemma 2. *Let $w_s : \mathbb{R} \mapsto \mathbb{R}_+$ be given by* (4). *Let $J_0 \in \mathbb{N}$. Hence, for $t > 0$, it holds that*

$$\sum_{j=-\infty}^{-J_0} \left|w_s\left(tB^{-2j}\right)\right|^2 = \frac{\chi_{s,B,J_0}(t)}{2\log B}\left(1 \pm O\left(|B-1|^2 \log|B-1|\right)\right), \tag{20}$$

where

$$\chi_{s,B,J_0}(t) := 2^{-2s}\Gamma\left(2s, 2tB^{\frac{J_0}{\log B}}\right).$$

Furthermore, it holds that

$$\sum_{j=J_0}^{\infty} \left|w_s\left(tB^{-2j}\right)\right|^2 = \frac{\phi_{s,B,J_0}(t)}{2\log B}\left(1 \pm O\left(|B-1|^2 \log|B-1|\right)\right), \tag{21}$$

where

$$\phi_{s,B,J_0}(t) := 2^{-2s}\gamma\left(2s, 2tB^{\frac{J_0}{\log B}}\right).$$

Proof. Let us start by proving (20). First of all, observe that the following identity holds:

$$\int_{B^{\frac{J_0}{\log B}}}^{\infty} |w_s(tx)|^2 \frac{dx}{x} = 2^{-2s}\Gamma\left(2s, 2tB^{\frac{J_0}{\log B}}\right) = \chi_{s,B,J_0}(t).$$

Applying an analogous procedure to the one adopted in Lemma 7.6 in [15], define the function $G_s : \mathbb{R} \mapsto \mathbb{R}_+$ by

$$G_s(u) := |w_s(e^u)|^2 = e^{-2e^u\left(1-sue^{-u}\right)}.$$

Let $j' = -j$, and fix $t = e^v$, $v > 0$; on one hand we get

$$\sum_{j'=J_0}^{\infty}\left|w_s\left(tB^{2j'}\right)\right|^2 = \sum_{j'=J_0}^{\infty}\left|w_s\left(e^{dj'+v}\right)\right|^2 = \sum_{j'=J_0}^{\infty} G_s\left(dj'+v\right),$$

$d = 2\log B$. On the other hand, for $u = \log x$, we obtain

$$\int_{B^{\frac{J_0}{\log B}}}^{\infty} |w_s(tx)|^2 \frac{dx}{x} = \int_{J_0}^{\infty}\left|w_s\left(e^{u+v}\right)\right|^2 du = \int_{J_0}^{\infty} G_s(u+v)\,du.$$

As in Lemma 7.6 in [15], note that $d\sum_{j'=J_0}^{\infty} G_s\left(dj'+v\right)$ is a Riemann sum for $\int_{B^{\frac{J_0}{\log B}}}^{\infty} G_s(u+v)\,du$. Moreover, because $\sum_{j'=J_0}^{\infty} G_s\left(dj'+v\right)$ is periodic with period d, it is sufficient to estimate this sum just for $0 < v < d$. Now observe that, for $J = J_0 + \Delta J$, $\Delta J > 0$, we get

$$\begin{aligned}
\left|d\sum_{j'=J_0}^{\infty} G_s\left(dj'+v\right) - \chi_{s,B,J_0}(t)\right| &= \left|d\sum_{j'=J_0}^{\infty} G_s\left(dj'+v\right) - \int_{J_0}^{\infty} G_s(u+v)\,du\right| \\
&\leq \left|d\sum_{j'=J_0}^{J} G_s\left(dj'+v\right) - \int_{J_0}^{Jd+\frac{d}{2}} G_s(u+v)du\right| \\
&\quad + d\sum_{j'>J} G_s\left(dj'+v\right) + \int_{Jd+\frac{d}{2}}^{\infty} G_s(u+v)\,du.
\end{aligned}$$

Using the midpoint rule, (see again Lemma 7.6 in [15]), we obtain

$$\left|d\sum_{j'=J_0}^{J} G_s\left(dj'+v\right) - \int_{J_0}^{Jd+\frac{d}{2}} G_s(u+v)\,du\right| \leq \frac{1}{24}\left\|G''\right\|_{\infty}(J-J_0)\,d^3.$$

On the other hand, observe that, for $r > 0$,

$$\frac{d}{dr}\left(1 - sre^{-r}\right) = se^{-r}(r-1),$$

so that $(1 - sre^{-r})$ is monotonically decreasing for $r \in [0,1)$, it attains its minimum for $r = 1$ and then it is monotonically increasing for $r \in (1,\infty)$. Observing that $(1 - sre^{-r})_{r=0} = 1$ and $\lim_{r\to\infty}(1 - sre^{-r}) = 1$ yields to

$$G_s(r) \leq e^{-2e^r\left(1-sre^{-r}\right)} \leq e^{-2e^r}.$$

Consequently, for $j'' = \exp d j'$, we obtain

$$d\sum_{j'>J} G_s\left(dj'+v\right) \leq \sum_{j'>J} e^{-2e^{dj'+v}} = \sum_{j''>\exp dJ} e^{-2e^v j''} = \frac{e^{2e^v\left(1-e^{dJ}\right)}}{e^{2e^v}-1},$$

which, for $y = 2e^{u+v}$, leads to

$$\begin{aligned}\int_{Jd+\frac{d}{2}}^{\infty} G_s(u+v)\,du &= \int_{Jd+\frac{d}{2}}^{\infty} e^{-2e^{u+v}}\,du \\ &= \int_{2e^{Jd+\frac{d}{2}+v}}^{\infty} e^{-y}\frac{dy}{y} \\ &\leq 2e^{-\left(Jd+\frac{d}{2}+v\right)}e^{-2e^{Jd+\frac{d}{2}+v}}.\end{aligned}$$

Therefore, there exists a constant $C > 0$ so that

$$\begin{aligned}\left|\sum_{j'=J_0}^{\infty} G_s\left(dj'+v\right) - \frac{\chi_{s,B,J_0}(t)}{d}\right| &\leq \frac{1}{24}\left\|G''\right\|_{\infty}\Delta J d^2 + \frac{e^{2e^v\left(1-e^{dJ}\right)}}{e^{2e^v}-1} \\ &\quad +2e^{-\left(Jd+\frac{d}{2}+v\right)}e^{-2e^{Jd+\frac{d}{2}+v}} \\ &\leq C\left(\Delta J d^2 + e^{-2e^{Jd}}\right) \leq C'\Delta J d^2.\end{aligned}$$

According again to [15], we choose $\Delta J \in (\log(1/d)/d, 2\log(1/d)/d)$ so that

$$\left|\sum_{j=-\infty}^{-J_0}\left|w_s\left(tB^{-2j}\right)\right|^2 - \frac{\chi_{s,B,J_0}(t)}{d}\right| \leq C'\left(2d\log\left(\frac{1}{d}\right)\right).$$

It follows that

$$\left|\left(\frac{\chi_{s,B,J_0}(t)}{d}\right)^{-1}\sum_{j=-\infty}^{-J_0}\left|w_s\left(tB^{-2j}\right)\right|^2 - 1\right| \leq 2\frac{C'}{\chi_{s,B,J_0}(t)}d^2\log\left(\frac{1}{d}\right).$$

Finally, because $d = 2\log B$ and $\lim_{B\to 1^+}\log B/(B-1) = 1$, the proof is complete. The proof of (21) is totally analogous and, therefore, omitted.

Lemma 3. *Let* $w_s : \mathbb{R} \mapsto \mathbb{R}_+$ *be given by* (4). *Then we have*

$$\sum_{|k|>K} w_s^2\left(\left(kB^{-j}\right)^2\right) \leq 2^{-\left(2s+\frac{1}{2}\right)}B^j\Gamma\left(2s+\frac{1}{2}, 2K^2B^{-2j}\right).$$

Proof. Observe that

$$\begin{aligned}\sum_{|k|>K} w_s^2\left(\left(kB^{-j}\right)^2\right) &= \sum_{|k|>K}\left(B^{-j}k\right)^{4s}\exp\left(-2\left(B^{-j}k\right)^2\right)\\ &\leq 2\int_K^\infty \left(B^{-j}x\right)^{4s}\exp\left(-2\left(B^{-j}x\right)^2\right)dx\\ &\leq 2^{-\left(2s+\frac{1}{2}\right)}B^j\left(\int_{2K^2B^{-2j}}^\infty u^{2s-\frac{1}{2}}\exp\left(-u\right)du\right)\\ &\leq 2^{-\left(2s+\frac{1}{2}\right)}B^j\Gamma\left(2s+\frac{1}{2},2K^2B^{2-j}\right),\end{aligned}$$

as claimed.

Corollary 1. *Let $\omega, J>0$; for x sufficiently large, it holds that*

$$\sum_{j=J_0}^{J} B^{-\omega j}\Gamma\left(S+1,xB^{-2j}\right)\leq C_{S,\alpha}x^S e^{-x}B^{-(\omega+2S)J}.$$

Proof. For x sufficiently large, the following limit holds

$$\lim_{x\to\infty}\frac{\Gamma\left(S+1,xB^{-2j}\right)}{\left(xB^{-2j'}\right)^S e^{-xB^{-2j}}}=1,$$

see for instance [1] Formula 6.5.32, pag. 263. Therefore, we have:

$$\sum_{j=0}^{J}B^{-(\omega+2S)j}x^S e^{-xB^{-2j}}\leq x^S e^{-x}\sum_{j=0}^{J}B^{-(\omega+2S)j}\leq C_{S,\omega}x^S e^{-x}B^{-(\omega+2S)J}.$$

The next result concerns the behaviour of the sums of the powers of the weights λ_{jq}.

Lemma 4. *Let Q_j and λ_{jq} be so that Theorem 1 holds. For any j, it holds that*

$$\sum_{q=1}^{Q_j}\lambda_{jq}\approx 1.$$

Furthermore, let $p>1$ and $j>0$. It holds that

$$\sum_{q=1}^{Q_j}\lambda_{jq}^p\leq \eta^{p-1}B^{j(1-p)}. \tag{22}$$

Proof. The first inequality follows directly the conditions in Theorem 1. On the other hand, for $j>0$, it is immediate to see that

$$\sum_{q=1}^{Q_j} \lambda_{jq}^p \leq \eta^p \sum_{q=1}^{Q_j} B^{-jp} \leq \eta^p Q_j B^{-jp} \leq \eta^{p-1} B^{j(1-p)}.$$

The next Lemma establishes explicit upper bounds for the sums with respect to q of differences between Mexican standard and truncated coefficients and for the L^2-norms of the sums with respect to q of Mexican standard and truncated needlets.

Lemma 5. *For $j > 0$, it holds that*

$$\sum_{q=1}^{Q_j} \left|\beta_{jq;sK} - \beta_{jq;s}\right|^2 \leq B^j \Gamma\left(2s+\frac{1}{2}, 2K^2B^{-2j}\right) \sum_{|k|>K} \gamma_k \, ; \tag{23}$$

$$\left\|\left(\psi_{jq;sK} - \psi_{jq;s}\right)\right\|^2_{L^2\left(\mathbb{S}^1\right)} \leq 2^{-\left(2s+\frac{1}{2}\right)} \eta \Gamma\left(2s+\frac{1}{2}, 2K^2B^{-2j}\right) . \tag{24}$$

Proof. For the Hölder inequality, it holds that

$$\begin{aligned}
\sum_{q=1}^{Q_j} \left|\beta_{jq;sK} - \beta_{jq;s}\right|^2 &= \sum_{q=1}^{Q_j} \lambda_{jq} \left|\sum_{|k|>K} w_s\left(\left(kB^{-j}\right)^2\right) a_k u_k\left(\xi_{jq}\right)\right|^2 \\
&\leq \sum_{q=1}^{Q_j} \lambda_{jq} \left(\sum_{|k|>K} w_s\left(\left(kB^{-j}\right)^2\right) \left|a_k\right| \left|u_k\left(\xi_{jq}\right)\right|\right)^2 \\
&\leq \sum_{q=1}^{Q_j} \lambda_{jq} \sum_{|k|>K} w_s^2\left(\left(kB^{-j}\right)^2\right) \sum_{|k|>K} \left|a_k\right|^2 \left|u_k\left(\xi_{jq}\right)\right|^2 \\
&\leq \sum_{q=1}^{Q_j} \lambda_{jq} \sum_{|k|>K} w_s^2\left(\left(kB^{-j}\right)^2\right) \sum_{|k|>K} \gamma_k.
\end{aligned}$$

Using Lemma 3 and Lemma 4 leads to the claimed result. As far as (24) is concerned, we get

$$\begin{aligned}
\left(\psi_{jq;sK}(\theta) - \psi_{jq;s}(\theta)\right)^2 &= \lambda_{jq} \left(\sum_{|k|>K} w_s\left(\left(kB^{-j}\right)^2\right) \overline{u_k\left(\xi_{jq}\right)} u_k(\theta)\right)^2 \\
&= \int_{\mathbb{S}^1} \lambda_{jq} \left(\sum_{|k_1|,|k_2|>K} w_s^2\left(\left(kB^{-j}\right)^2\right)\right).
\end{aligned}$$

Using the orthogonality of $\{u_k\}$ and Lemma 3 yields to

$$\left\|\left(\psi_{jq;sK}-\psi_{jq;s}\right)\right\|_{L^2\left(\mathbb{S}^1\right)}^2 = \int_{\mathbb{S}^1} \lambda_{jq}\left(\sum_{|k_1|,|k_2|>K} w_s^2\left(\left(kB^{-j}\right)^2\right)\right)$$
$$\leq 2^{-\left(2s+\frac{1}{2}\right)}\eta\Gamma\left(2s+\frac{1}{2},2K^2B^{-2j}\right),$$

as claimed.

6.2 Ancillary results related to Theorem 3

The lemmas here proved describe the behaviour of I_1, I_2 and I_3. Hence, they are pivotal to study the bias $R_{s,K,J,\eta}$.

Lemma 6. *Let I_1 be given by*

$$I_1 := \left\|\sum_{j>J}\sum_{q=1}^{Q_j}\beta_{jq;s}\psi_{jq;s}\right\|_{L^2\left(\mathbb{S}^1\right)}.$$

Then, there exists $C_1>0$ such that

$$I_1 \leq C_1 B^{-rJ}.$$

Proof. Note preliminarily that

$$I_1 \leq \sum_{j>J}\left\|\sum_{q=1}^{Q_j}\beta_{jq;s}\psi_{jq;s}\right\|_{L^2\left(\mathbb{S}^1\right)}.$$

Observe that, for the Hölder inequality (see also [3]), we get

$$\left(\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\psi_{jq;s}(\theta)\right|\right)^2 = \left(\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\right|\left|\psi_{jq;s}(\theta)\right|^{\frac{1}{2}}\left|\psi_{jq;s}(\theta)\right|^{\frac{1}{2}}\right)^2$$
$$\leq \left(\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\right|^2\left|\psi_{jq;s}(\theta)\right|\right)\left(\sum_{q=1}^{Q_j}\left|\psi_{jq;s}(\theta)\right|\right)$$
$$\leq C\eta^{\frac{1}{2}}B^{\frac{j}{2}}\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\right|^2\left|\psi_{jq;s}(\theta)\right|.$$

For $C>0$ and using (10), it follows that

$$\begin{aligned}\left\|\sum_{q=1}^{Q_j}\beta_{jq;s}\psi_{jq;s}\right\|^2_{L^2(\mathbb{S}^1)} &\leq C\eta^{\frac{1}{2}}B^{\frac{j}{2}}\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\right|^2\left\|\psi_{jq;s}\right\|_{L^1(\mathbb{S}^1)}\\ &\leq C\sum_{q=1}^{Q_j}\eta\left|\beta_{jq;s}\right|^2 \leq CB^{-2rj}.\end{aligned}$$

Hence, we obtain

$$I_{1,1}=\sum_{j>J}\left\|\sum_{q=1}^{Q_j}\beta_{jq;s}\psi_{jq;s}\right\|_{L^2(\mathbb{S}^1)} \leq C_{1,2}B^{-rJ}\ .$$

Lemma 7. *Let* I_2 *be given by*

$$I_2:=\left\|\sum_{j=J_0}^{J}\sum_{q=1}^{Q_j}\beta_{jq;s}\left(\psi_{jq;s}-\psi_{jq;sK}\right)\right\|_{L^2(\mathbb{S}^1)}.$$

Then, there exists $C_2>0$ *such that*

$$I_2\leq C_2J^{\frac{1}{2}}K^{2s-\frac{1}{2}}\exp\left(-K^2\right)B^{-\left(r+2s-\frac{1}{2}\right)J}.$$

Proof. First of all, observe

$$\begin{aligned}&\left\|\sum_{j=J_0}^{J}\sum_{q=1}^{Q_j}\beta_{jq;s}\left(\psi_{jq;s}-\psi_{jq;sK}\right)\right\|^2_{L^2(\mathbb{S}^1)}\\ &\qquad\leq (J-J_0+1)\sum_{j=J_0}^{J}\left\|\sum_{q=1}^{Q_j}\beta_{jq;s}\left(\psi_{jq;s}-\psi_{jq;sK}\right)\right\|^2_{L^2(\mathbb{S}^1)};\end{aligned}$$

using the Hölder inequality, we have

$$\begin{aligned}\left(\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\left(\psi_{jq;s}(\theta)-\psi_{jq;sK}(\theta)\right)\right|\right)^2 \leq &\left(\sum_{q=1}^{Q_j}\left(\psi_{jq;s}(\theta)-\psi_{jq;sK}(\theta)\right)^2\right)\\ &\left(\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\right|^2\right),\end{aligned}$$

so that

$$\left\|\sum_{q=1}^{Q_j}\beta_{jq;s}\left(\psi_{jq;s}-\psi_{jq;sK}\right)\right\|^2_{L^2(\mathbb{S}^1)}\leq\left(\sum_{q=1}^{Q_j}\left|\beta_{jq;s}\right|^2\right)\sum_{q=1}^{Q_j}\left\|\left(\psi_{jq;sK}-\psi_{jq;s}\right)\right\|^2_{L^2(\mathbb{S}^1)}.$$

Using (24) in Lemma 5 and (10), we obtain

$$\left\| \sum_{q=1}^{Q_j} \beta_{jq;s} \left(\psi_{jq;s} - \psi_{jq;sK} \right) \right\|_{L^2\left(\mathbb{S}^1\right)}^2 \leq CB^{-2rj} \Gamma \left(2s + \frac{1}{2}, 2K^2 B^{-2j} \right)$$

so that, from the Corollary 1, it follows

$$\sum_{j=J_0}^{J} B^{-2rj} \Gamma \left(2s + \frac{1}{2}, 2K^2 B^{-2j} \right) \leq CK^{4s-1} \exp\left(-2K^2\right) B^{-(2r+4s-1)J}.$$

Hence, we get

$$I_2 \leq C_2 J^{\frac{1}{2}} K^{2s-\frac{1}{2}} \exp\left(-K^2\right) B^{-\left(r+2s-\frac{1}{2}\right)J},$$

as claimed.

Lemma 8. *Let I_3 be given by*

$$I_3 := \left\| \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \left(\beta_{jq;s} - \beta_{jq;sK} \right) \psi_{jq;sK} \right\|_{L^2\left(\mathbb{S}^1\right)} .$$

Then, there exists $C_3 > 0$ such that

$$I_3 \leq C_3 B^{\left(\frac{3}{2}-2s\right)J} J^{\frac{1}{2}} K^{2s-\frac{1}{2}} e^{-K^2} \left(\sum_{|k|>K} \gamma_k \right)^{\frac{1}{2}} .$$

Proof. Observe that

$$\left\| \sum_{j=J_0}^{J} \sum_{q=1}^{Q_j} \left(\beta_{jq;s} - \beta_{jq;sK} \right) \psi_{jq;sK} \right\|_{L^2\left(\mathbb{S}^1\right)}^2 \leq (J - J_0 + 1) \sum_{j=J_0}^{J} \left\| \sum_{q=1}^{Q_j} \left(\beta_{jq;s} - \beta_{jq;sK} \right) \psi_{jq;sK} \right\|_{L^2\left(\mathbb{S}^1\right)}^2 .$$

Hölder inequality leads to

$$\left(\sum_{q=1}^{Q_j} \left| \left(\beta_{jq;s} - \beta_{jq;sK} \right) \psi_{jq;sK}(\theta) \right| \right)^2 \leq \left(\sum_{q=1}^{Q_j} \left(\beta_{jq;s} - \beta_{jq;sK} \right)^2 \left| \psi_{jq;sK}(\theta) \right| \right) \left(\sum_{q=1}^{Q_j} \left| \psi_{jq;sK}(\theta) \right| \right).$$

We therefore obtain

$$\sum_{q=1}^{Q_j} \left| \psi_{jq;s,K}(\theta) \right| \leq \sum_{q=1}^{Q_j} \left| \psi_{jq;s}(\theta) \right| \leq CB^{\frac{j}{2}};$$

so that

$$\left\| \sum_{q=1}^{Q_j} \left(\beta_{jq;s} - \beta_{jq;sK}\right) \psi_{jq;sK} \right\|^2_{L^2(\mathbb{S}^1)} \leq CB^{\frac{j}{2}} \sum_{q=1}^{Q_j} \left(\beta_{jq;s} - \beta_{jq;sK}\right)^2 \left\| \psi_{jq;sK} \right\|_{L^1(\mathbb{S}^1)}$$
$$\leq CB^j \Gamma\left(2s + \frac{1}{2}, 2K^2 B^{-2j}\right) \sum_{|k|>K} \gamma_k.$$

Using Corollary 1 leads to

$$\sum_{j=J_0}^{J} B^j \Gamma\left(2s + \frac{1}{2}, 2K^2 B^{-2j}\right) \leq C\left(2K^2\right)^{2s-\frac{1}{2}} \exp\left(-2K^2\right) B^{-2(2s-1)J}.$$

Hence, we get

$$I_3 \leq C_3 B^{(1-2s)J} J^{\frac{1}{2}} K^{s-\frac{1}{4}} e^{-2K^2} \left(\sum_{|k|>K} \gamma_k\right)^{\frac{1}{2}},$$

as claimed.

6.3 Ancillary results related to Theorem 2

In this subsection we will summon auxiliary results connected to the proof Theorem 2.

Lemma 9. *Let $E_{1,1}$, $E_{1,2}$, $E_{1,3}$ and $E_{1,4}$ be given respectively by* (16), (17), (18) *and* (19). *Then, there exists $C_E > 0$ such that*

$$E_{1,1} + E_{1,2} + E_{1,3} + E_{1,4} \leq C_E \left(\frac{n}{\log n}\right)^{-\frac{2r}{2r+1}}.$$

Proof. Observe that

$$E_{1,1} \leq C_1 \eta_n n^{-1} \left(\sum_{j=0}^{J_n} \sum_{q=1}^{Q_j} \mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right| \geq \frac{\kappa\tau_n}{2}\right\}}\right),$$

Splitting the sum into two parts by means of the so-called optimal bandwidth selection, given by $J_{1,n} : B^{J_{1,n}} = (n/\log n)^{\frac{1}{2r+1}}$, and using (26) yields to

$$\eta_n \sum_{j=0}^{J_{1,n}} \sum_{q=1}^{Q_j} \mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right| \geq \frac{\kappa\tau_n}{2}\right\}} \leq CB^{J_{1,n}} \leq C\left(n/\log n\right)^{\frac{1}{2r+1}},$$

and, on the other hand,

$$\eta_n \sum_{j=J_{1,n}}^{J_n} \mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right| \geq \frac{\kappa\tau_n}{2}\right\}} \leq C\eta_n \sum_{j=J_{1,n}}^{J_n} \sum_{q=1}^{Q_j} \left|\beta_{jq;sK_n}\right|^2 \left(\frac{\kappa\tau_n}{2}\right)^{-2}$$
$$\leq C' \frac{n}{\log n} B^{-2rJ_{1n}}$$
$$\leq C' \left(\frac{n}{\log n}\right)^{\frac{1}{2r+1}}.$$

It follows

$$E_{1,1} \leq C_{1,1} \left(\frac{n}{\log n}\right)^{-\frac{2r}{2r+1}}.$$

As far as $E_{1,2}$ is concerned, standard calculations using (27) lead to

$$E_{1,2} = \eta_n \sum_{j=0}^{J_n} \sum_{q=1}^{Q_j} \mathbb{E}\left[\left|\widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n}\right|^2 \mathbb{1}_{\left\{\left|\widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n}\right| \geq \kappa\tau_n/2\right\}}\right]$$
$$\leq C\eta_n \sum_{j=0}^{J_n} \sum_{q=1}^{Q_j} \mathbb{E}^{\frac{1}{2}}\left[\left|\widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n}\right|^4\right] \mathbb{P}^{\frac{1}{2}}\left[\left|\widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n}\right| \geq \frac{\kappa\tau_n}{2}\right]$$
$$\leq C' \sum_{j=0}^{J_n} B^j n^{-1} n^{-\frac{\delta}{2}} \leq C'' B^{J_n} n^{-1} n^{-\frac{\delta}{2}} \leq C_{1,2}\eta_n (\log n)^{-1} n^{-\frac{\delta}{2}}.$$

On the other hand, we obtain

$$E_{1,3} = \eta_n \sum_{j=0}^{J_n} \sum_{q=1}^{Q_j} \left|\beta_{jq;sK_n}\right|^2 \mathbb{E}\left[\mathbb{1}_{\left\{\left|\widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n}\right| \geq \kappa\tau_n/2\right\}}\right]$$
$$\leq C_{1,3} n^{-\delta} \|F\|^2_{L^2(\mathbb{S}^1)}.$$

Finally, we get

$$E_{1,4} \leq C\eta_n \sum_{j=0}^{J_n} \sum_{q=1}^{Q_j} \left|\beta_{jq;sK_n}\right|^2 \mathbb{1}_{\left\{\left|\beta_{jq;sK_n}\right| < 2\kappa\tau_n\right\}}$$
$$\leq C \sum_{j=0}^{J_{1,n}} \sum_{q=1}^{Q_j} |2\kappa\tau_n|^2 + \eta_n \sum_{j=J_{1,n}}^{J_n} \sum_{q=1}^{Q_j} \left|\beta_{jq;sK_n}\right|^2$$
$$\leq C' \left(B^{J_{1,n}} \left(\frac{n}{\log n}\right)^{-1} + \sum_{j=J_{1,n}}^{J_n} B^{-2rj}\right),$$

so that

$$E_{1,4} \leq C_{1,4} \left(\frac{n}{\log n}\right)^{-\frac{2r}{2r+1}},$$

as claimed.

The next result was originally presented in [3] as Lemma 16, hence the proof is here omitted.

Lemma 10. *Let σ be a finite positive constant such that*

$$\sigma \geq \left(\|F\|_{L^{\infty}\left(\mathbb{S}^1\right)} \left\| \psi_{jq;s} \right\|^2_{L^2\left(\mathbb{S}^1\right)} \right)^{\frac{1}{2}} .$$

Then, there exists constants $c_P, c_E, C > 0$ such that, for $B^j \leq \left(\frac{n}{\log n}\right)^{\frac{1}{2}}$, the following inequalities hold

$$\mathbb{P}\left[\left| \widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n} \right| > x \right] \leq 2 \exp\left(\frac{nx^2}{2\left(\sigma^2 + c_P x B^{\frac{j}{2}}\right)} \right); \tag{25}$$

$$\mathbb{E}\left[\left| \widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n} \right|^2 \right] \leq c_E n^{-1}; \tag{26}$$

$$\mathbb{P}\left[\left| \widehat{\beta}_{jq;sK_n} - \beta_{jq;sK_n} \right| > \frac{\kappa \tau_n}{2} \right] \leq C n^{-\delta} \tag{27}$$

where $\delta \geq 6\sigma^2$.

Acknowledgements - The author whishes to thank D. Marinucci and I.Z. Pesenson for the useful suggestions and discussions.

References

1. **Abramowitz, M. and Stegun, I. (1946).***Handbook of mathematical functions.* Dover, New York.
2. **Al-Sharadqah, A., and Chernov, N. (2009).** Error analysis for circle fitting algorithms. *Electron. J. Stat., 3, 886–911.*
3. **Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D. (2009).** Adaptive density estimation for directional data using Needlets. *Ann. Statist., 37 (6A), 3362-3395.*
4. **Bhattacharya, A. and Bhattacharya, R. (2008).** Nonparametric statistics on manifolds with applications to shape spaces. *IMS Lecture Series.*
5. **Di Marzio, M., Panzera, A. and Taylor, C.C. (2009).** Local polynomial regression for circular predictors. *Stat. Probab. Lett., 79 (19), 2066–2075.*
6. **Donoho, D. and Johnstone, I. (1994).** Ideal spatial adaptation via wavelet shrinkage. *Biomefrika, 81, 425-455.*
7. **Donoho, D., Johnstone, I., Kerkyacharian, G. and Picard, D. (1996).** Density estimation by wavelet thresholding. *Ann. Statist., 24, 508–539.*
8. **Durastanti, C., Geller, D. and Marinucci, D. (2012).** Adaptive nonparametric regression on spin fiber bundles. *J. Multivariate Anal., 104 (1), 16–38.*
9. **Durastanti, C. and Lan, X., (2013).** High-frequency tail index estimation by nearly tight frames. *AMS Contemporary Mathematics, Vol. 603.*
10. **Durastanti, C. (2013).** Tail behaviour of Mexican needlets. *Submitted.*
11. **Durastanti, C. (2015).** Block Thresholding on the sphere. *Sankhya A, 77 (1), 153–185.*
12. **Durastanti, C. (2016).** Quantitative central limit theorems for Mexican needlet coefficients on circular Poisson fields. *to appear on Stat. Methods Appl.*

13. **Durastanti, C. (2016).** Adaptive global thresholding on the sphere. *Submitted.*
14. **Fisher, N.I. (1993).** *Statistical Analysis of Circular Data*. Cambridge University Press.
15. **Geller, D. and Mayeli, A. (2006).** Continuous wavelets and frames on stratified Lie groups I. *J. Fourier Anal. Appl., 12 (5), 543–579.*
16. **Geller, D. and Mayeli, A. (2009).** Continuous wavelets on manifolds. *Math. Z., 262, 895–927*
17. **Geller, D. and Mayeli, A. (2009).** Nearly tight frames and space-frequency analysis on compact manifolds. *Math. Z., 263, 235–264.*
18. **Geller, D. and Mayeli, A. (2009).** Besov spaces and frames on compact manifolds. *Indiana Univ. Math. J., 58, 2003–2042*
19. **Geller, D. and Pesenson, I.Z. (2011).** Band-limited localized Parseval frames and Besov spaces on compact homogeneous manifolds. *J. Geom. Anal., 21 (2), 334–371.*
20. **Hardle, W., Kerkyacharian, G., Picard, D., and Tsybakov, A. (1998)**. *Wavelets, approximation and statistical applications.* Springer.
21. **Kato, S., Shimizu, K. and Shieh, G. S. (2008).** A circular–circular regression model. *Statist. Sinica, 18 (2), 633–645.*
22. **Klemela, J. (2000).** Estimation of densities and derivatives of densities with directional data. *J. Multivariate Anal., 73, 18–40.*
23. **Lan, X. and Marinucci, D. (2009).** On the dependence structure of wavelet coefficients for spherical random fields. *Stochastic Process. Appl., 119, 3749-3766.*
24. **Marinucci, D. and Peccati, G. (2011).***Random fields on the sphere.* Cambridge University Press.
25. **Mayeli, A. (2010).** Asymptotic uncorrelation for Mexican needlets. *J. Math. Anal. Appl., 363 (1), 336–344.*
26. **Narcowich, F.J., Petrushev, P. and Ward, J.D. (2006a).** Localized tight frames on spheres. *SIAM J. Math. Anal., 38, 574–594.*
27. **Narcowich, F.J., Petrushev, P. and Ward, J.D. (2006b).** Decomposition of Besov and Triebel-Lizorkin spaces on the sphere. *J. Funct. Anal, 238 (2), 530–564.*
28. **Pesenson, I.Z. (2013)**, Multiresolution analysis on compact Riemannian manifolds. *in Multiscale analysis and nonlinear dynamics, Rev. Nonlinear Dyn. Complex, Wiley-VCH, 65-82.*
29. **Rao Jammalamadaka, S. and Sengupta, A. (2001).** *Topics in circular statistics.* World Scientific.
30. **Scodeller, S., Rudjord, O. Hansen, F.K., Marinucci, D., Geller, D. and Mayeli, A. (2011).** Introducing Mexican needlets for CMB analysis: issues for practical applications and comparison with standard needlets. *ApJ, 733 (121).*
31. **Silverman, B.W. (1986).** *Density estimation for statistics and data analysis.* Chapman & Hall CRC.
32. **Stein, E. and Weiss, G. (1971).** *Introduction to Fourier analysis on Euclidean spaces.* Princeton University Press.
33. **Tsybakov, A.B. (2009).** *Introduction to Nonparametric estimation*, Springer, New York.
34. **Wu, H. (1997).** Optimal exact designs on a circle or a circular arc. *Ann. Statist., 25(5), 2027–2043.*